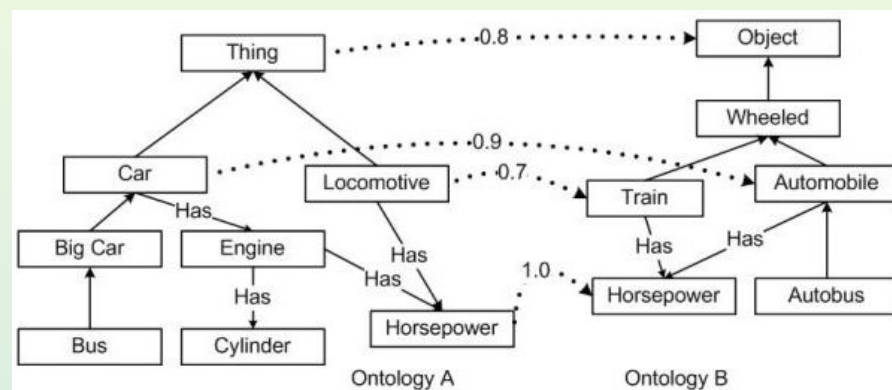


SYMMETry

Semantic similarity Matching Multiple ontologies

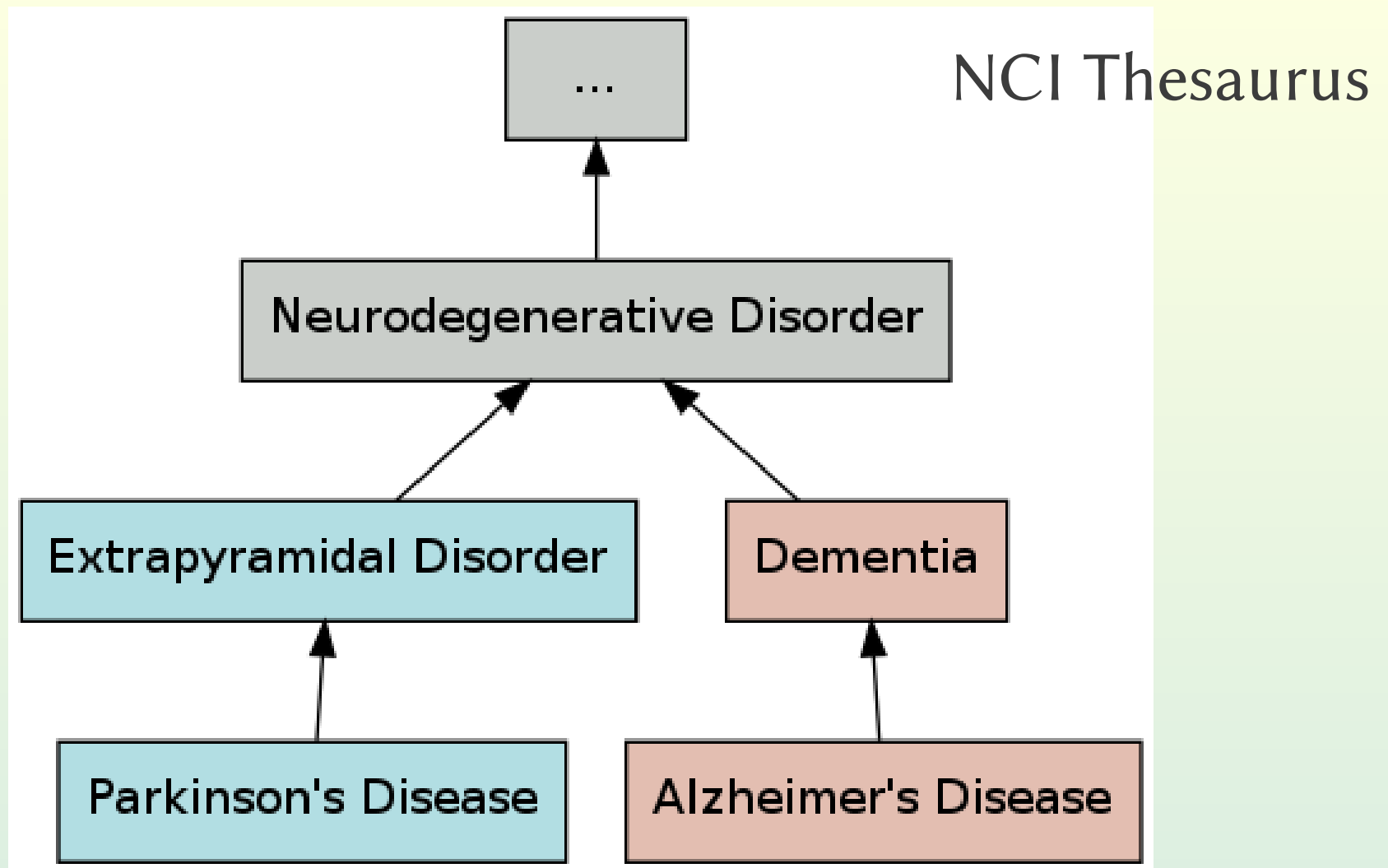
João D Ferreira
Catia Pesquita
Francisco M Couto



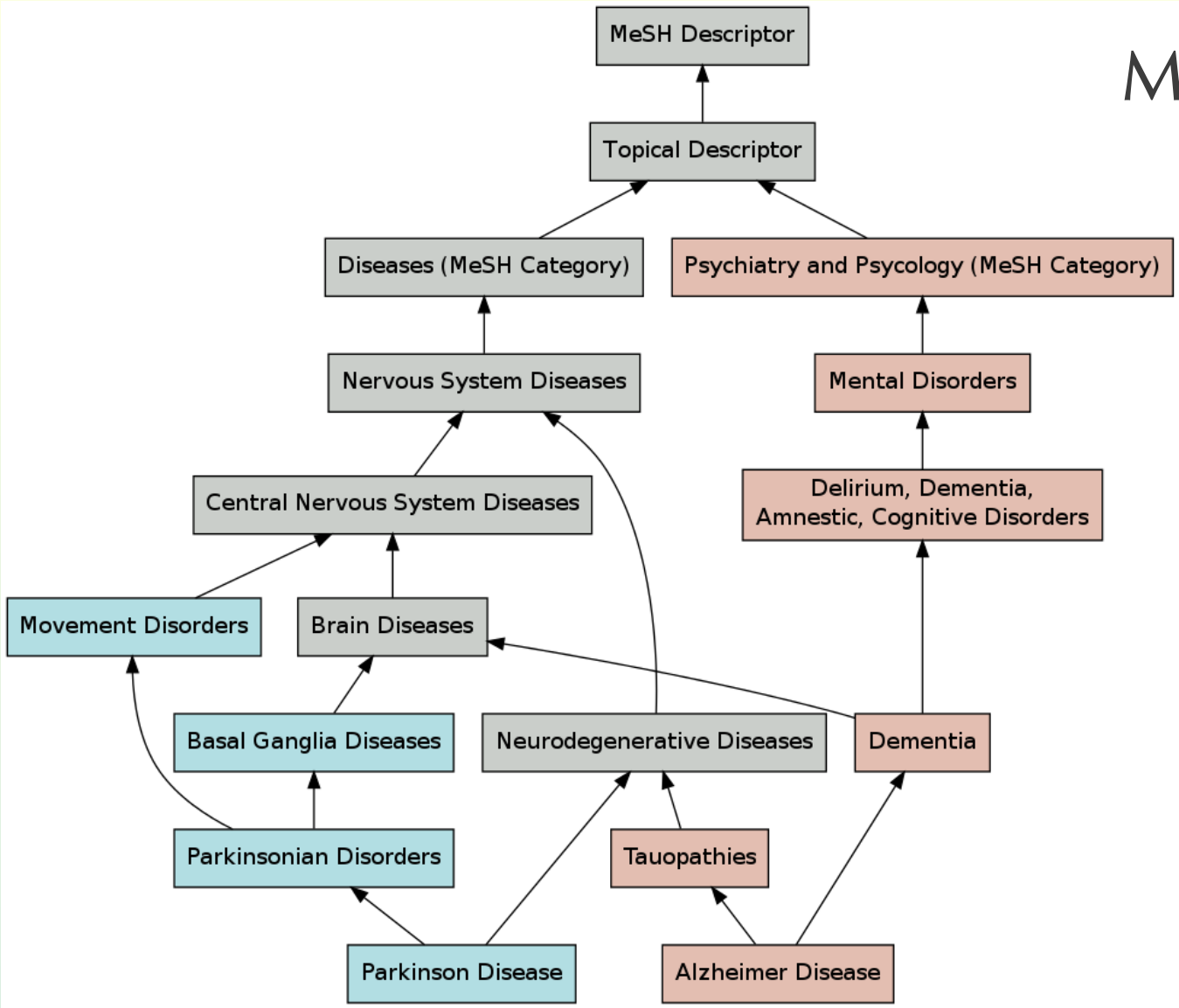
Motivation

- Main objective: compare biomedical entities
 - The concepts between themselves
 - Complex entities annotated with concepts from multiple ontologies
- Let's look at an example:
 - Parkinson's Disease \leftrightarrow Alzheimer's Disease

Motivation



Motivation



MeSH

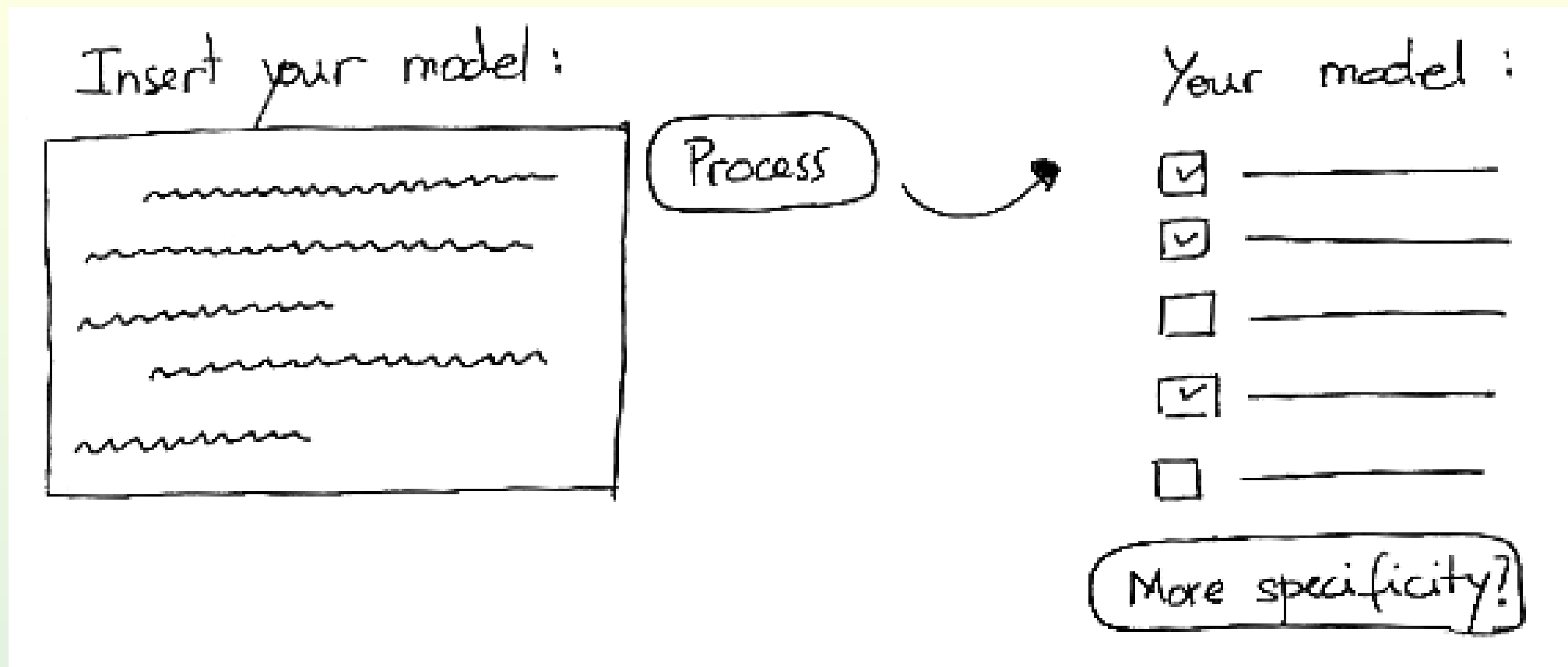
Motivation

- Parkinson's Disease:
 - Risks → rural environment, pesticides
 - Protective factors → smoking
 - Symptoms → tremor, rigidity, postural instability, cognitive and behavioral problems
 - Causes → insufficient dopamine, alpha-synuclein
 - Treatment → dopamine agonists, MAO-B inhibitors
 - Associated genes → SNCA, UCH-L1, PRKN

Motivation

- Alzheimer's Disease:
 - Risks → mutation on APOE4
 - Diagnosis → behavioral assessment, cognitive tests, brain scans
 - Symptoms → confusion, aggression, mood swings, long-term memory loss, bodily functions loss
 - Causes → plaques in brain, tangles in brain
 - Treatment → *unknown*
 - Preventions → mental stimulation
 - Associated proteins → beta-amyloid

Mock-up



Mock-up

Two models :



















Compare

Mock-up

Result :

SSM

_____	x	_____	0.91
_____	x	_____	0.83
_____	x	{ _____ _____	0.59

Measures

- sim GIC
- Resnik
- BMA

Total : 0.71

Mock-up

Insert your model:



 Process



Your model:



More specificity?

Two models:

- ☒ _____
- ☒ _____
- ☒ _____
- ☒ _____
- ☒ _____

- ☒ _____
- ☒ _____
- ☒ _____
- ☒ _____
- ☒ _____
- ☒ _____
- ☒ _____
- ☒ _____
- ☒ _____

Compare »

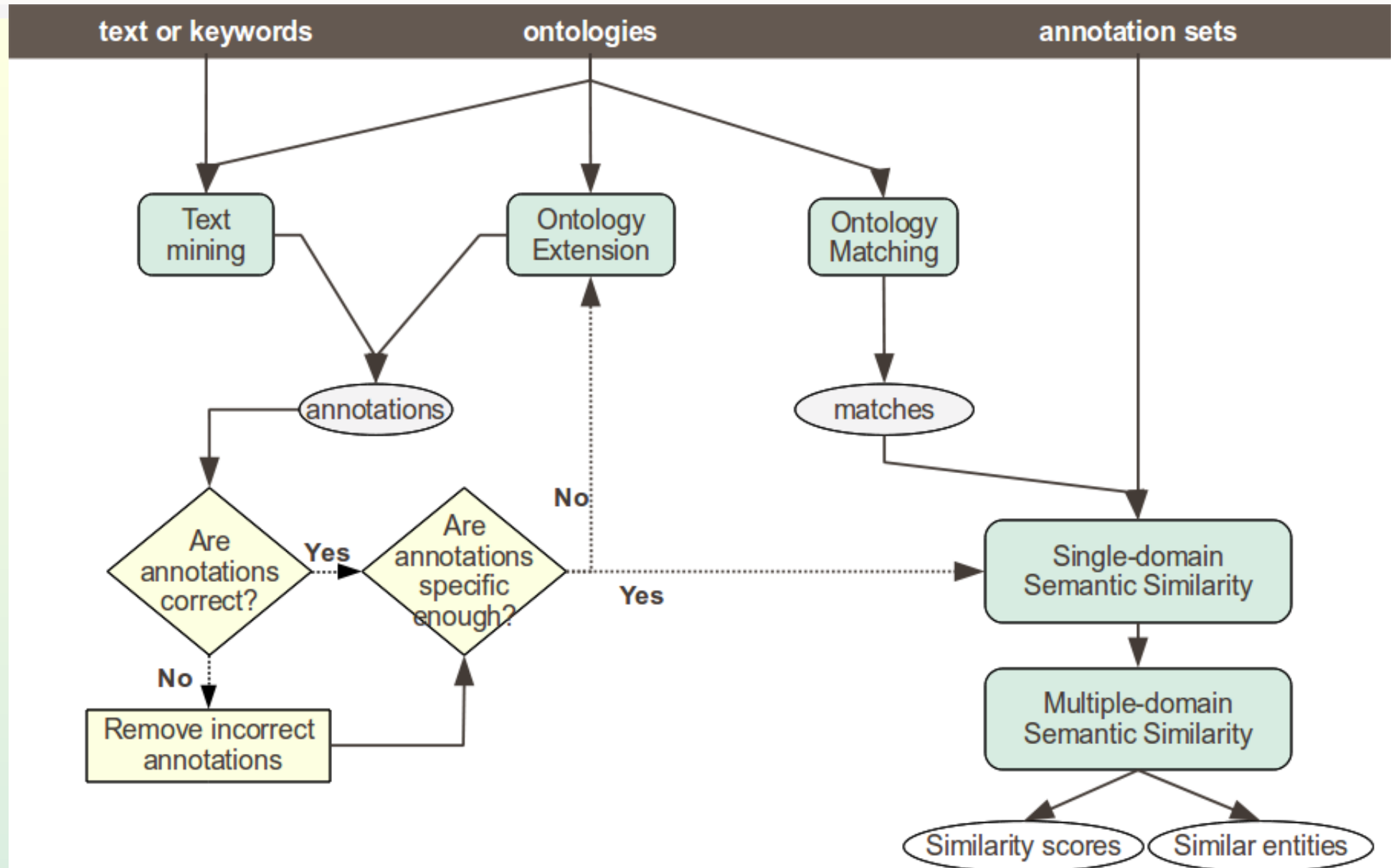
Mock-up

Results:

	SSM	Measure
_____ x _____	0.91	simGIC
_____ x _____	0.83	Resnik
_____ x { _____ _____	0.59	
...		

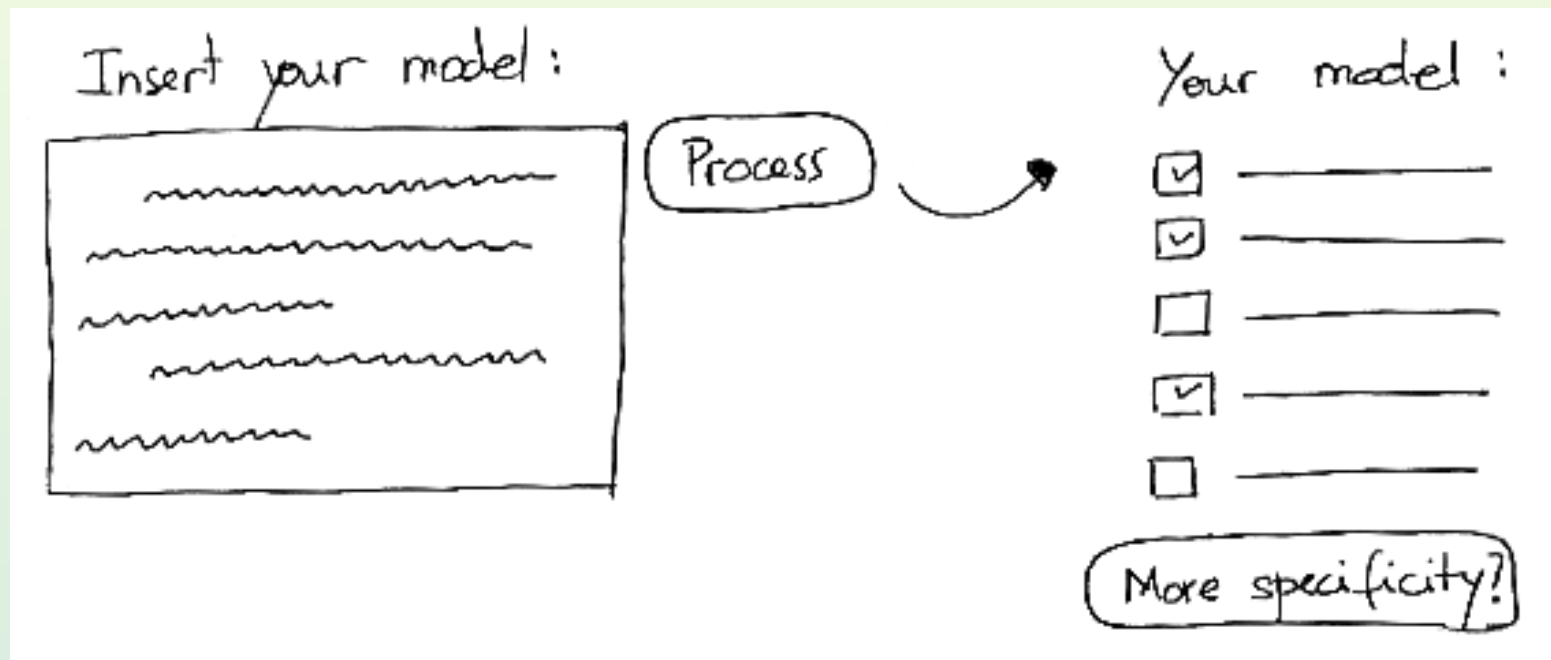
Total: 0.73

Methodology



TEXT
MININGONTOLOGY
EXTENSIONONTOLOGY
MATCHINGSINGLE-
DOMAIN
SSMMULTI-
DOMAIN
SSM

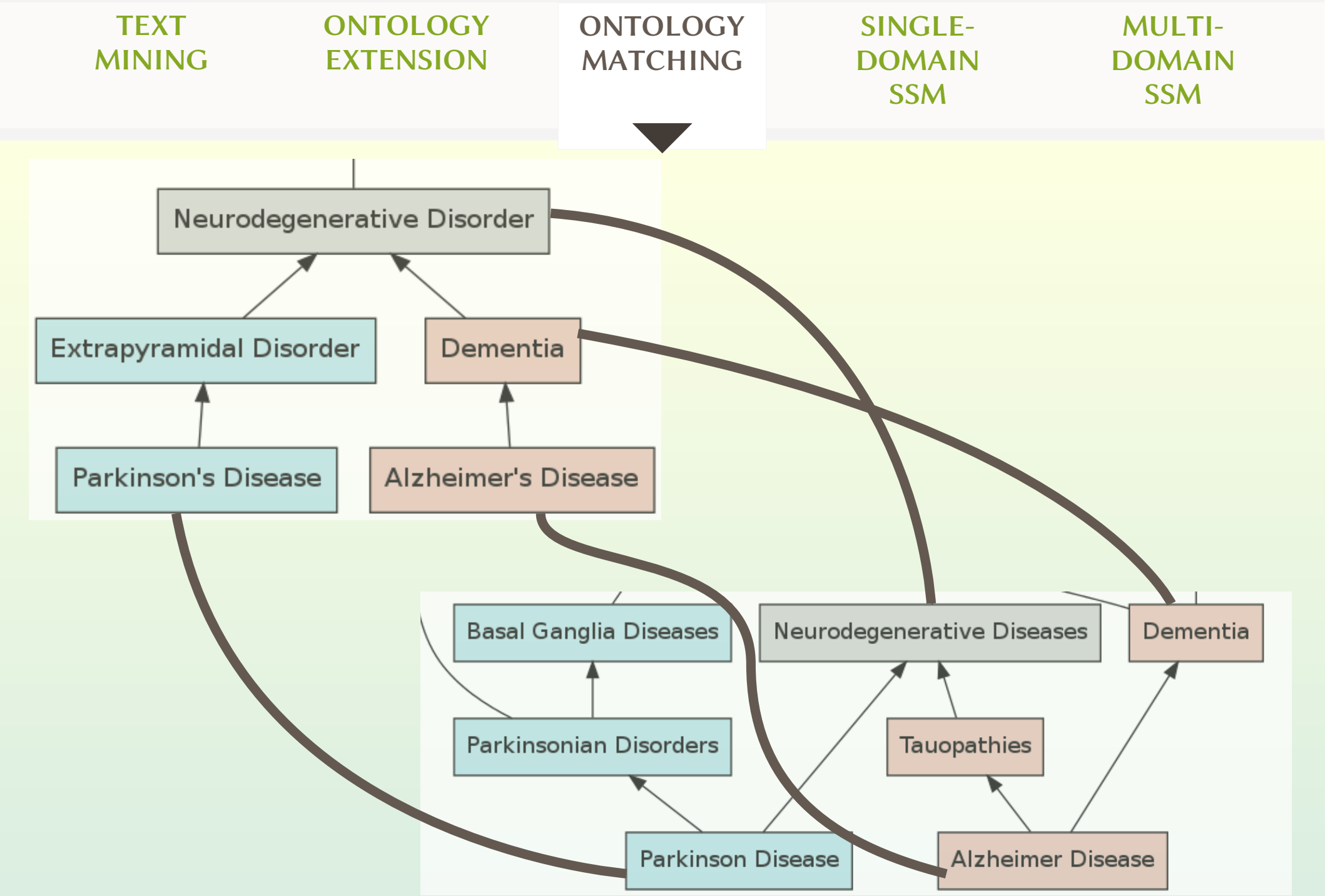
- **Text Mining & Ontology Extension:**
 - Assist the user in describing their models with concepts from existing ontologies
 - Ontology Matching also helps the process



TEXT
MININGONTOLOGY
EXTENSIONONTOLOGY
MATCHINGSINGLE-
DOMAIN
SSMMULTI-
DOMAIN
SSM

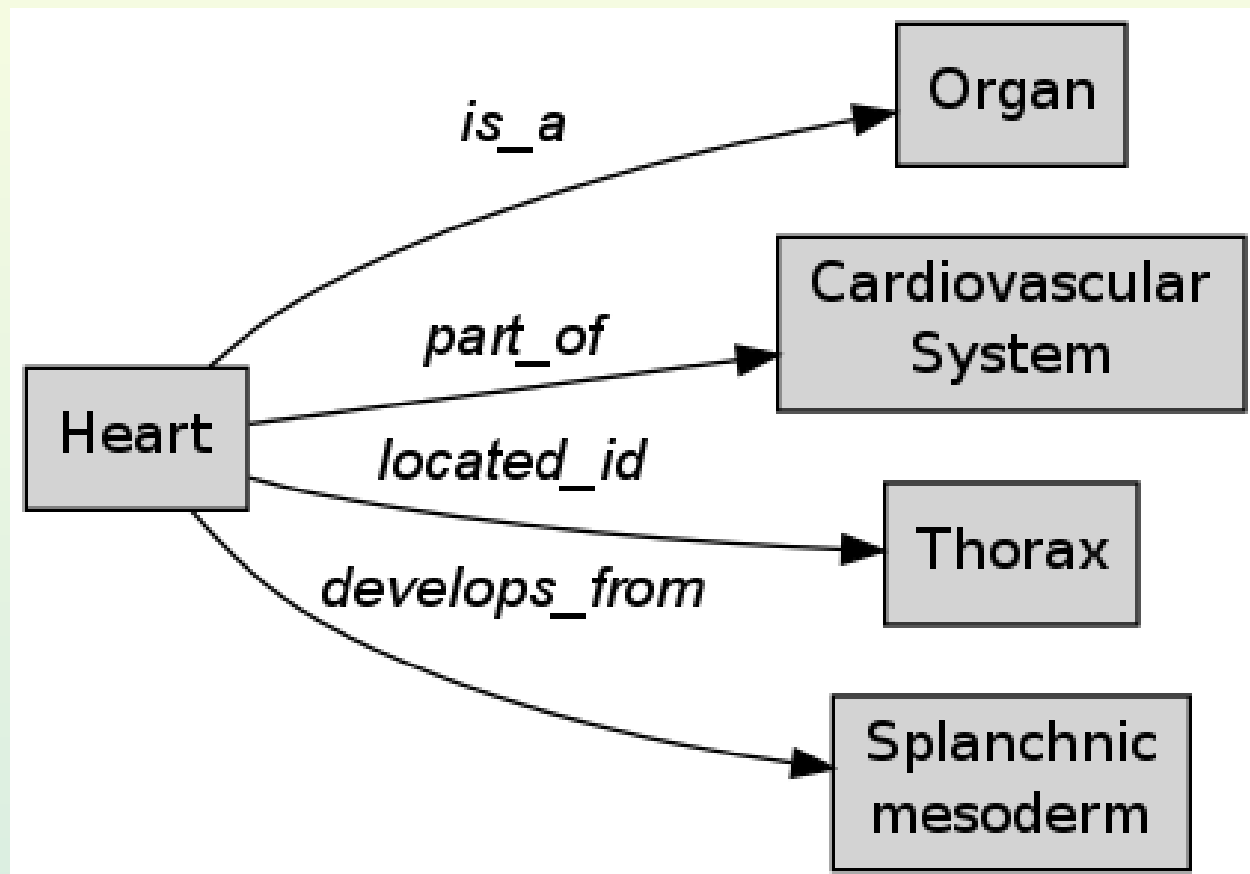
■ Ontology Matching:

- Two concepts from different ontologies are matched
- Links of *equivalence*, *broader* and *narrower* scope, *relatedness*, ...



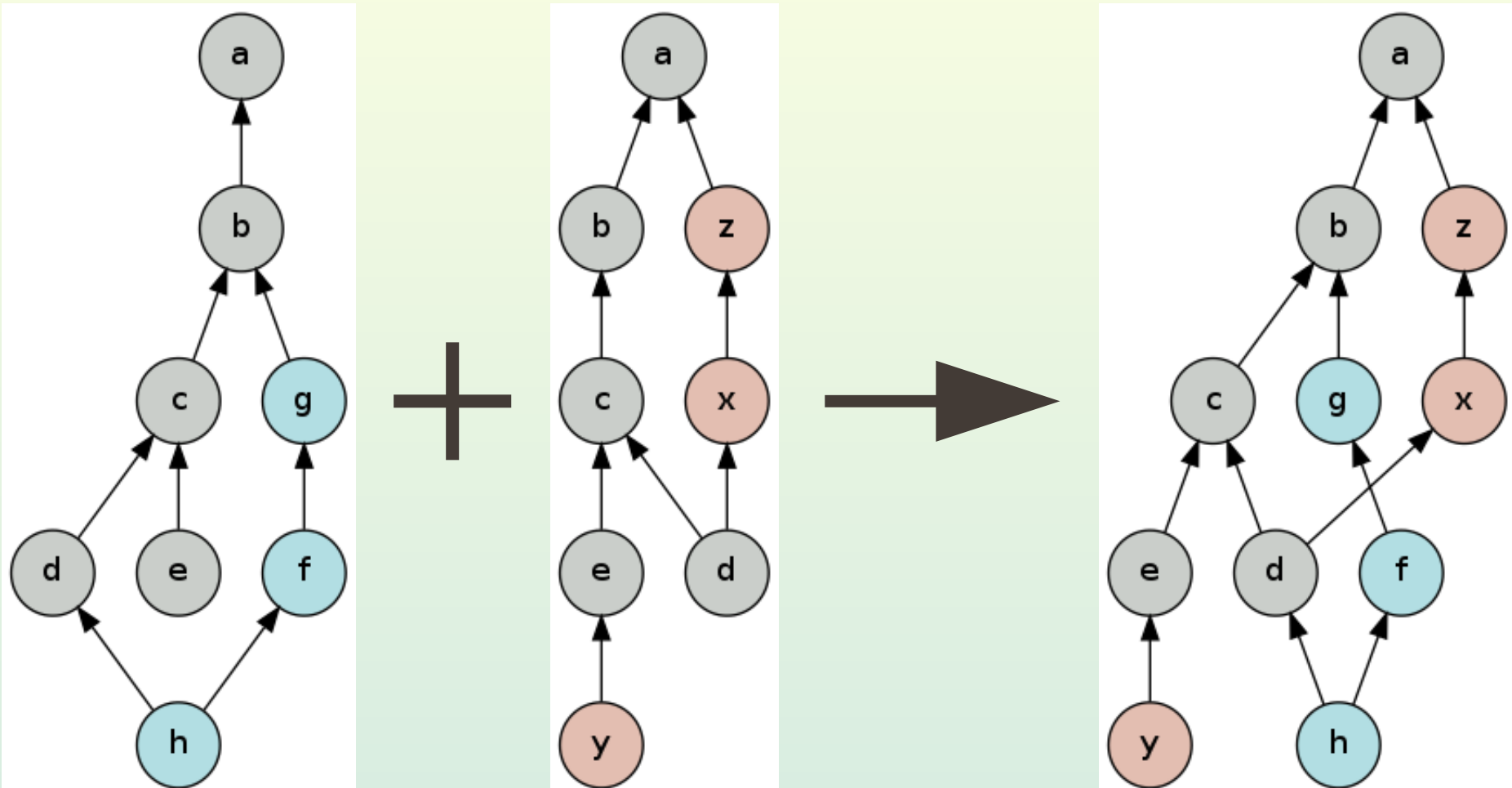
TEXT
MININGONTOLOGY
EXTENSIONONTOLOGY
MATCHINGSINGLE-
DOMAIN
SSM
▼MULTI-
DOMAIN
SSM

- **Single-domain semantic similarity**
- How to define ancestry:
 - *is_a* ?
 - *part_of* ?
 - *located_id* ?
 - *develops_from* ?



TEXT
MININGONTOLOGY
EXTENSIONONTOLOGY
MATCHINGSINGLE-
DOMAIN
SSM
▼MULTI-
DOMAIN
SSM

- IC (measure of specificity): intrinsic/extrinsic
- Based on single ontology or on multiple ontologies

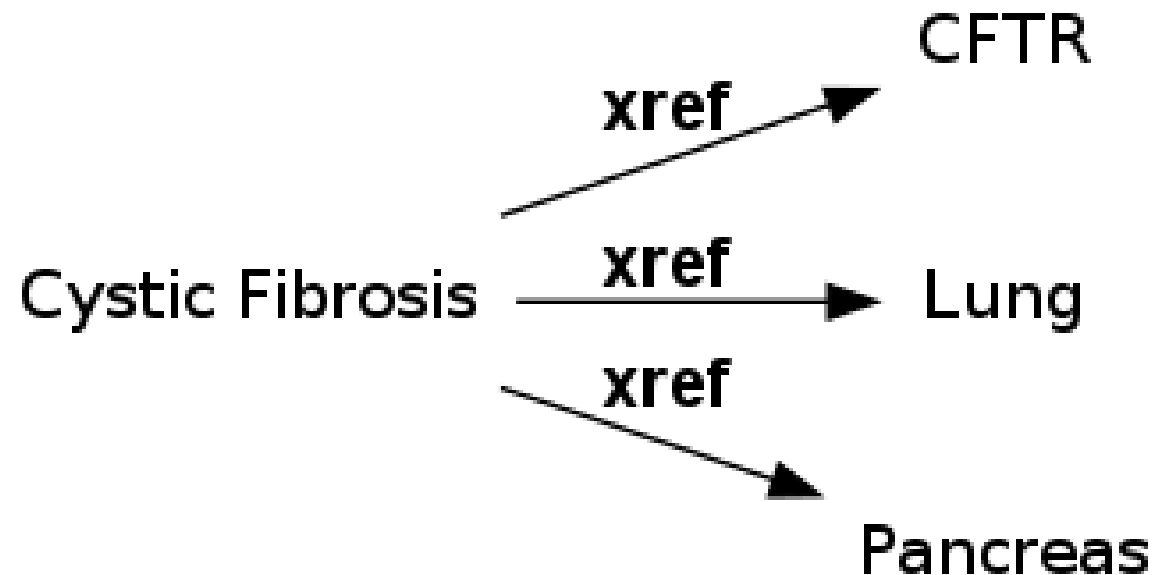


TEXT
MININGONTOLOGY
EXTENSIONONTOLOGY
MATCHINGSINGLE-
DOMAIN
SSMMULTI-
DOMAIN
SSM

- Common schema of relations (Relation Ontology?)
- Edge-based methods (incorporate matches between concepts from different ontologies)
- Comparison of entities annotated with single-domain concepts

TEXT
MININGONTOLOGY
EXTENSIONONTOLOGY
MATCHINGSINGLE-
DOMAIN
SSMMULTI-
DOMAIN
SSM

- Multi-domain semantic similarity
- Cross-references contribute to SSM of concepts:
 - *CFTR* and *Cystic Fibrosis*
 - *CFTR* and *Lung*
 - ...



TEXT
MININGONTOLOGY
EXTENSIONONTOLOGY
MATCHINGSINGLE-
DOMAIN
SSMMULTI-
DOMAIN
SSM

- Parkinson's Disease:
 - Risks → rural environment, pesticides
 - Protective factors → smoking
 - Symptoms → tremor, rigidity, postural instability, cognitive and behavioral problems
 - Causes → insufficient dopamine, alpha-synuclein
 - Treatment → dopamine agonists, MAO-B inhibitors
 - Associated genes → SNCA, UCH-L1, PRKN

TEXT
MININGONTOLOGY
EXTENSIONONTOLOGY
MATCHINGSINGLE-
DOMAIN
SSMMULTI-
DOMAIN
SSM

- Alzheimer's Disease:
 - Risks → mutation on APOE4
 - Diagnosis → behavioral assessment, cognitive tests, brain scans
 - Symptoms → confusion, aggression, mood swings, long-term memory loss, bodily functions loss
 - Causes → plaques in brain, tangles in brain
 - Treatment → *unknown*
 - Preventions → mental stimulation
 - Associated proteins → beta-amyloid

TEXT
MININGONTOLOGY
EXTENSIONONTOLOGY
MATCHINGSINGLE-
DOMAIN
SSMMULTI-
DOMAIN
SSM

- Distinct interests
 - Physicians want risk factors, treatments, ...
 - Other researchers want causes, genes, proteins, ...

⇒ **Weights!**

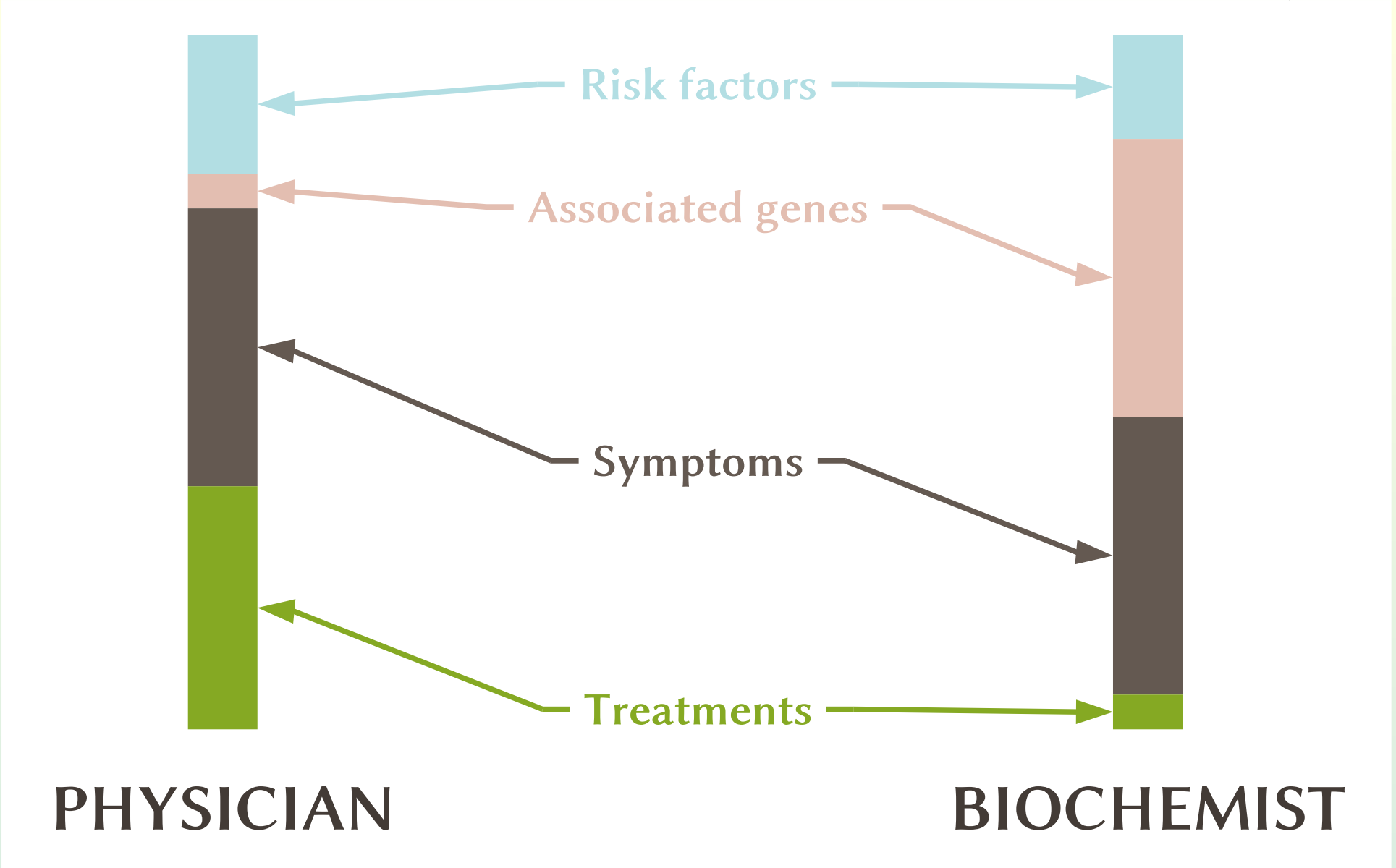
TEXT
MINING

ONTOLOGY
EXTENSION

ONTOLOGY
MATCHING

SINGLE-
DOMAIN
SSM

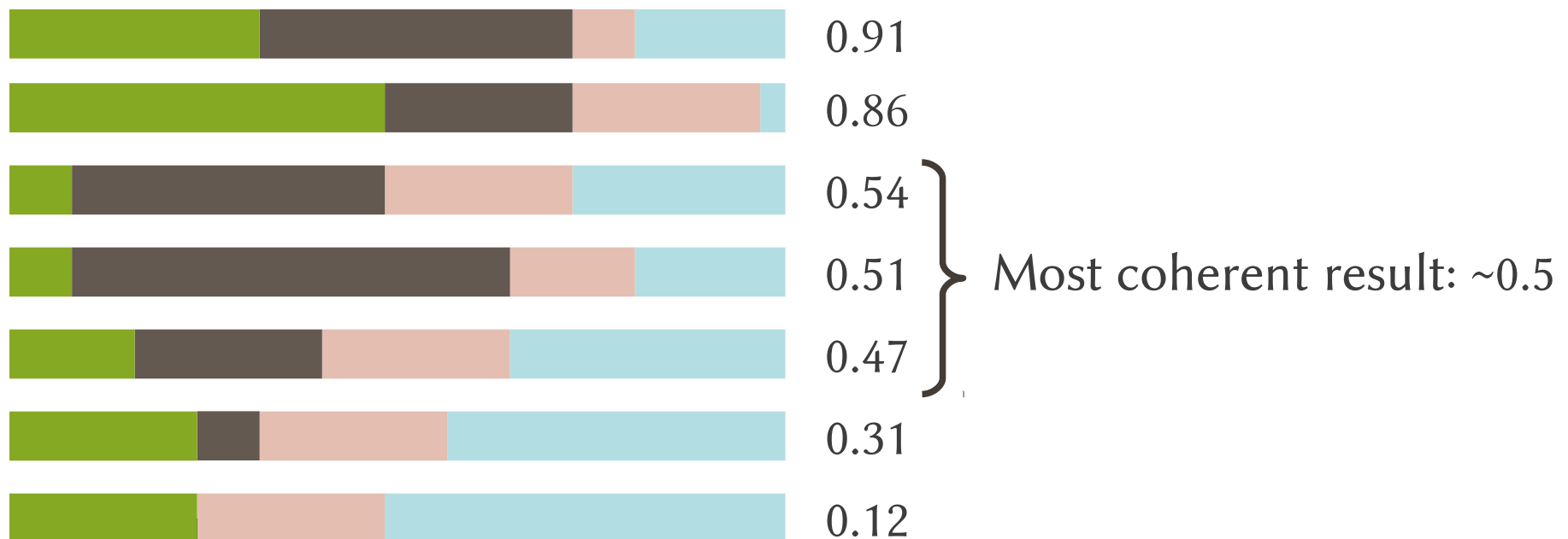
MULTI-
DOMAIN
SSM



TEXT
MININGONTOLOGY
EXTENSIONONTOLOGY
MATCHINGSINGLE-
DOMAIN
SSMMULTI-
DOMAIN
SSM

- The user chooses a specific configuration;
- Or all configurations, choosing a coherent score

⇒ Voting strategy



BioXLDB · LaSIGE · FCUL

SYMMETry

**Semantic similarity Matching
Multiple ontologies**